Smoothing Methods for Nonlinear Complementarity Problems

Mounir Haddou, Patrick Maheux

Noname manuscript No.
(will be inserted by the editor)

Smoothing Methods for Nonlinear Complementarity Problems

Mounir Haddou · Patrick Maheux

Received: date / Accepted: date

Abstract In this paper, we present some new smoothing techniques to solve general nonlinear complementarity problems. Under a weaker condition than monotonicity on the original problems, we prove convergence of our methods. We also present an error estimate under a general monotonicity condition. Some numerical tests confirm the efficiency of the proposed methods.

Keywords Complementarity problem · Smoothing function · Error estimate · Trajectory.

Mathematics Subject Classification (2000) 90C33.

1 Introduction

The class of nonlinear complementarity problems (NCP) is a simple and popular tool for addressing practical problems arising in mathematical programming, economics, engineering, and the sciences (see for instance [1] and [2]). For example, general equilibrium modeling, trajectory problems, traffic network design, mechanical contact problems, obstacle prob-

Communicated by Jean-Pierre Crouzeix

Mounir Haddou (✉)
INSA-IRMAR, UMR-CNRS 6625, Rennes, France.
E-mail: mounir.haddou@insa-rennes.fr

Patrick Maheux
MAPMO, UMR-CNRS 7349, Fédération Denis Poisson, Université d'Orléans, France
E-mail: patrick.maheux@univ-orleans.fr

lems, bimatrix games, etc., lead naturally to the solution of NCP. Over the last years, the numerical optimization community has devoted considerable energies toward developing efficient algorithms and methods for solving these problems and there exists a considerable body of literature presenting algorithms for NCP.

Although the effectiveness of complementarity algorithms has improved substantially in recent years, the fact remains that increasingly more difficult problems are being proposed that are exceeding the capabilities of these algorithms. As a result, there is a real need to propose new methods and algorithms to address complicated and difficult situations.

In particular, a large part of existing algorithms concerns linear complementarity problems and NCP satisfying particular conditions (monotonicity, ...). Our purpose, in this paper, is to propose a class of simple and efficient methods for solving nonlinear complementarity problems even in the complicated cases, when a weaker condition than monotonicity is used.

The paper is organized as follows. In Section 2, we present the considered problems and precise the assumptions. In Section 3, we define the smoothing functions and the approximation technique to solve general nonlinear complementarity problem. We also propose a generic way to construct such functions. In Section 4, we discuss in details the properties of the smoothing functions and the approximation scheme. This section is also devoted to the proof of convergence and error estimates. A generic algorithm and numerical examples and experiments are presented in Section 5 and a conclusion and some perspectives are presented at the end of the paper.

2 Preliminaries and Problem Setting

Consider the nonlinear complementarity problem NCP, which is to find a solution of the system :

$$x \geq 0,\ F(x) \geq 0 \quad \text{and} \quad x^\top F(x) = 0, \tag{1}$$

where $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a continuous function satisfying some additional assumptions to be precised later. NCP is equivalent to find $\bar{x} \geq 0$ satisfying the following variational inequality

$$\langle F(\bar{x}), x - \bar{x} \rangle \geq 0, \qquad \forall x \geq 0.$$

To solve NCP, there are essentially three different classes of methods: equation-based methods (smoothing), merit functions and projection-type methods.

Our goal in this paper is to present new and very simple smoothing and approximation schemes to solve NCP and to produce efficient numerical methods; see Theorem 3.3 and Theorem 4.1.

First, let us introduce usual assumptions on F and the ones that will be used in this paper. A well known and studied situation corresponds to monotone functions F and several methods and algorithms have been developed in this case. We recall that F is said to be monotone iff $F : \mathbb{R}^n \to \mathbb{R}^n$ satisfies for any $x \neq y \in \mathbb{R}^n$,

$$(x-y)^\top (F(x) - F(y)) \geq 0. \tag{2}$$

In our work, we will consider P_0-functions, this will include a larger class of functions. A function $F : \mathbb{R}^n \to \mathbb{R}^n$ is said to be a P_0-function iff

$$\max_{i: x_i \neq y_i} (x-y)_i (F_i(x) - F_i(y)) \geq 0, \quad \forall\, x \neq y \in \mathbb{R}^n,$$

and, when this inequality is strict, F is said to be a P-function.

The P_0-assumption is obviously weaker than monotonicity. The notion of P_0-function is a generalization of the notion of P_0-matrices that are matrices $M \in \mathbb{R}^{n \times n}$ with non-negative principal minors. This concept is extensively used in linear complementary problem, i.e. when F and M are related by an equation of the form $F(x) = Mx + q$ for some $q \in \mathbb{R}^n$; see [3, p.153]. This notion is also well-known and used in nonlinear complementary problem; see for instance [4].

Note that, when F is monotone, the set of solutions of NCP is convex (possibly empty). When F is a P-function, the set of solutions (if non empty) is a singleton. Neither of these

two properties are true under the condition P_0. Indeed, for instance,
$F(x,y) = (1-x, y-1), (x,y) \in \mathbb{R}^2$ is a P_0-function and the solution set of the corresponding NCP is $S_F = \{(0,1),(1,1)\}$ which is nonconvex and nonconnected.
We assume throughout that the solution set $\mathscr{X} := \{x \geq 0, F(x) \geq 0,\ x^\top F(x) = 0\}$ of (NCP) is nonempty and bounded.
An example of sufficient condition on the mapping F that insures the boundedness of $\mathscr{X}$ is provided by the following lemma.

Lemma 2.1 *Let F be a continuous and monotone function defined from $\mathbb{R}^{\mathrm{n}}$ into itself. Assume that there exist $y \in \mathbb{R}^{\mathrm{n}}$ and two positive constants M, c such that*

(i) $F(y) > 0$ *and* $c\,||y|| < \min_i F_i(y)$, *and*

(ii) *for any x,* $||x||_1 \geq M \Longrightarrow ||F(x)|| \leq c\,||x||_1$.

Then $\mathscr{X}_\varepsilon := \{x \in \mathbb{R}^{\mathrm{n}} : \mathrm{x} \geq 0, \mathrm{F(x)} \geq 0, \mathrm{x}^\top \mathrm{F(x)} \leq \varepsilon\}$ *is bounded for any* $\varepsilon \geq 0$.

Proof For convenience, set $m(F(y)) := \min_i F_i(y)$. Since F is continuous, the set $\mathscr{X}_\varepsilon$ is closed. The monotonicity property of F implies for any $x \in \mathscr{X}_\varepsilon$:

$$x^\top F(y) \leq x^\top F(x) - y^\top F(x) + y^\top F(y) \leq ||y||.||F(x)|| + ||y||.||F(y)|| + \varepsilon.$$

Let $x \in \mathscr{X}_\varepsilon$ and $||x||_1 \geq M$. Then

$$m(F(y)).||x||_1 \leq \sum_{i=1}^{n} x_i F_i(y) \leq c||y||.||x||_1 + ||y||.||F(y)|| + \varepsilon.$$

Thus, $\quad (m(F(y)) - c\,||y||).||x||_1 \leq ||y||.||F(y)|| + \varepsilon.$
Since $\kappa := m(F(y)) - c\,||y|| > 0$, we get $\quad ||x||_1 \leq (||y||.||F(y)|| + \varepsilon)/\kappa.$
Hence $\quad ||x||_1 \leq \max(M, (||y||.||F(y)|| + \varepsilon)/\kappa)$. So, $\mathscr{X}_\varepsilon$ is bounded. $\square$

All bounded continuous monotone functions F satisfying the first part of condition (i), i.e. $(F(y) > 0)$ also satisfy part 2 of condition (i) and condition (ii) of Lemma 2.1. Indeed there exists $R > 0$ such that for any $M > 0$ and any x with $||x||_1 \geq M$,

$$||F(x)|| \leq R \leq \frac{R}{M}||x||_1.$$

For M large enough and $c := \frac{R}{M}$, we obtain $c\,||y|| < m(F(y))$.

The condition (ii) allows us to consider functions F with sublinear growth at infinity.

A necessary and sufficient condition for boundedness and non emptiness of the set of solutions is known in the more general case of pseudomonotone variational inequality problems [5, Theorem 2 and Corollary 3]. There are also several other existence results when F is a P-function (a classical one is the existence of a strictly interior point i.e. $y > 0$ such that $F(y) > 0$).

We define componentwise the function $F_{\min}$ by $F_{\min}(x) := (F_{\min,i}(x))_{i:1...n}$ with $F_{\min,i}(x) = \min(x_i, F_i(x))$ for any $i : 1...n$. This function will play an important role in our study. Indeed, any solution of the equation $F_{min}(x) = 0$ is a also a solution of the (NCP) and conversely. Furthermore, F_{min} preserves some of the properties of F.

Lemma 2.2 *Assume that F is a P_0 (respectively P)-function then F_{min} is also P_0 (respectively P)-function.*

Proof Assume that F is a P_0-function. For any $x \neq y \in \mathbb{R}^{\mathrm{n}}$, there exists i such that $x_i \neq y_i$ and $(x_i - y_i)(F_i(x) - F_i(y)) \geq 0$. Without loss of generality, we can assume $x_i > y_i$. So, we get $F_i(x) \geq F_i(y)$. Now since $F_i(x) \geq \min(y_i, F_i(y))$ and $x_i \geq \min(y_i, F_i(y))$, it follows that $\min(x_i, F_i(x)) \geq \min(y_i, F_i(y))$.

Finally $(x_i - y_i)(F_{min,i}(x) - F_{min,i}(y)) \geq 0$ i.e. F_{min} is a P_0-function. The proof is analogue if F is a P-function. □

3 The Smoothing Functions

In the first part of this section, we introduce the smoothing functions and establish different properties that will be useful for the presentation and the convergence of the algorithm. The second part of this section proposes a generic way to construct such smoothing functions.

3.1 Definition and properties of the smoothing functions

We start our discussion by introducing the function θ with the following properties (These functions were used in [6,7]). Let $\theta : \mathbb{R} \to]-\infty, 1[$ be a non-decreasing continuous function such that

$$\theta(t) < 0 \text{ if } t < 0, \quad \theta(0) = 0 \quad \text{and} \lim_{t \to +\infty} \theta(t) = 1.$$

For instance, $\theta^{(1)}(t) = \frac{t}{t+1}$, $t \geq 0$ and $\theta^{(1)}(t) = t$ if $t < 0$, $\theta^{(2)}(t) = 1 - e^{-t}$, $t \in \mathbb{R}$. We will often return to these two examples very different from each other. We will also use these two functions in the numerical section.

In order to "detect" if $t = 0$ or $t > 0$ in a "continuous way", we introduce $\theta_r(t) := \theta(t/r)$ for $r > 0$. Indeed, $\theta_r(0) = 0$ for all $r > 0$ and $\lim_{r \searrow 0} \theta_r(t) = 1$ for all $t > 0$. Now, let us discuss the following equation in the one-dimensional case. Let $s, t \in \mathbb{R}^+$, be such that

$$\theta_r(s) + \theta_r(t) = 1. \tag{3}$$

For instance, let us take $\theta^{(1)}$. The equality (3) is then equivalent to $st = r^2$.
So, when r goes to 0, we simply get $st = 0$. The equation (3) applied with $s = x \in \mathbb{R}^+$ and $t = F(x) \in \mathbb{R}^+$ is then an approximation of the relation $xF(x) = 0$.

Our aim is to propose a large class of θ's for which the problems

$$x^{(r)} \geq 0, \; F(x^{(r)}) \geq 0 \quad \text{and} \quad \theta_r(x^{(r)}) + \theta_r(F(x^{(r)})) = 1 \tag{4}$$

are well posed and any limit point of $(x^{(r)})$ when, r goes to 0, is a solution of (NCP).

In the multidimensional case, the equation just above has to be interpreted as a system of n equations,

$$\theta_r(x_i^{(r)}) + \theta_r(F_i(x^{(r)})) = 1, \qquad i : 1...n.$$

Note that the relation (4) is symmetric in x and $F(x)$. Thus, our problem can be seen as a fixed point problem for the function $F_{r,\theta}(x)$ defined just below. Indeed, the equation (4) is equivalent to

$$x = \theta_r^{-1}\left(1 - \theta_r(F(x))\right) = r\theta^{-1}\left(1 - \theta(F(x)/r)\right) =: F_{r,\theta}(x).$$

By symmetry of the equation (3), we also have the relations:

$$F(x) = \theta_r^{-1}(1-\theta_r(x)) = r\theta^{-1}(1-\theta(x/r)).$$

But we shall not go that direction. We propose another way to approximate a solution of the (NCP) problem as follows. Let $\psi_r(t) = 1-\theta_r(t)$; the relation (4) is equivalent to the three following equalities

$$\psi_r(x)+\psi_r(F(x)) = 1 = \psi_r(0),$$
$$\psi_r^{-1}[\psi_r(x)+\psi_r(F(x))] = 0 \text{ and}$$
$$r\psi^{-1}\left[\psi\left(\tfrac{x}{r}\right)+\psi\left(\tfrac{F(x)}{r}\right)\right] = 0.$$

(with $\psi = \psi_1 = 1-\theta_1$). For the sequel, we set for any $x, y \in \mathbb{R}^n$ and any $r > 0$

$$G_r(x,y) := r\psi^{-1}\left[\psi\left(\frac{x}{r}\right)+\psi\left(\frac{y}{r}\right)\right]. \tag{5}$$

First, we characterize the solutions (x,y) of $G_r(x,y) = 0$ when ψ satisfies some conditions independent of F.

Let $0 < a < 1$. We say that ψ satisfies (H_a) if there exists $s_a > 0$ such that, for all $s \geq s_a$,

$$\psi(s) \leq \frac{1}{2}\psi(as) \quad \text{or equivalently} \quad \frac{1}{2}+\frac{1}{2}\theta(as) \leq \theta(s). \tag{H_a}$$

The condition (H_a) imposes that the decay of $\psi(s)$ is under some uniform control for large s or in terms of θ that $\theta(s)$ should grow enough quickly with some uniformity for large s. Since ψ and θ are monotone, it is interesting to take a as large as possible in the condition (H_a) since $(H_a) \Longrightarrow (H_b)$ for $b < a$.

Note that we can never take $a = 1$ because $\theta \leq 1$ unless θ is constant and equal to one for large s. But in some cases, a can be chosen as close to 1, see for instance $\theta^{(2)}$ and Theorem 3.2 below.

Note that this condition is satisfied for both "extreme" examples and many other examples.

1. For $\theta^{(1)}$, we have $\psi^{(1)}(t) = \frac{1}{t+1}$ if $t \geq 0$ and $\psi^{(1)}(t) = 1-t$ if $t < 0$ and the condition (H_a) is only satisfied for $0 < a < 1/2$ with $s_a \geq \frac{1}{1-2a}$.

2. For $\theta^{(2)}$, we have $\psi^{(2)}(t) = e^{-t}$ and the condition (H_a) is satisfied for any $0 < a < 1$ with $s_a = \frac{\ln 2}{1-a}$.

Note that these functions ψ do not satisfy the condition (H_a) in the same range for a. This has some consequence for the limit of $G_r(s,t)$ as r goes to zero (see Theorem 3.2).

Remark 3.1 *The condition (H_a) is not so restrictive. In fact there are plenty of functions satisfying this condition. For instance, if ψ satisfies (H_a) for a given $a \in]0,1[$ then ψ^β also satisfies this condition for any $\beta \geq 1$. More generically, a large family can be built on ψ by composition. Let $G :]0,+\infty[\longrightarrow]0,+\infty[$ be a nondecreasing function satisfying the conditions below,*

$$
\begin{aligned}
&(i)\ \ G(1) = 1, \quad \lim_{u\to 0^+} G(u) = 0,\\
&(ii)\ \lim_{u\to+\infty} G(u) = +\infty,\\
&(iii)\ G(u/2) \leq \tfrac{1}{2} G(u),
\end{aligned}
$$

where $0 < u < u_0$ for some $u_0 > 0$. Then $G \circ \psi$ satisfies (H_a) for the same a. For instance $G_1(u) = u^\beta e^{c(u^\alpha - 1)}$ and $G_2(u) = u^\beta \left(\log(eu^\alpha)\right)^\gamma$ satisfy $(i)-(ii)$ for any $\beta \geq 1$ and any $c, \alpha, \gamma \geq 0$ with $u_0 = +\infty$.

Other functions as $\theta(t) = 1 - e^{-t^\alpha}$, $t > 0$ with $\alpha > 0$ and as $\theta(t) = \frac{e^t - 1}{e^t + 1}$, $t > 0$ satisfy (H_a) for any $a \in]0,1[$.

From now on, all the results use the function ψ. Obviously, everything can be easily transposed on θ. We shall need the following lemma in Theorem 3.1 and Theorem 3.3

Lemma 3.1 *If $\psi : \mathbb{R} \longrightarrow]0,+\infty[$ is an invertible non-increasing function, then for any $s,t,r > 0$, we have $G_r(s,t) \leq \min(s,t)$. (Where G_r defined by (5))*

Proof Let $s,t \in \mathbb{R}$ be fixed. By symmetry, we can assume that $s = \min(s,t)$.
Since $\psi \geq 0$, we obviously have $\quad \psi(s/r) \leq \psi(s/r) + \psi(t/r)$.
By the fact that ψ is invertible and non-increasing, we get

$$\psi^{-1}\left(\psi(s/r) + \psi(t/r)\right) \leq s/r.$$

Thus, from the definition of G_r we conclude that

$$G_r(s,t) = r\psi^{-1}\left[\psi\left(s/r\right) + \psi\left(t/r\right)\right] \leq s = \min(s,t). \qquad \square$$

The next theorem shows how the condition (H_a) gives information about the behavior of G_r.

Theorem 3.1 *Let $\psi : \mathbb{R} \longrightarrow]0,+\infty[$ be an invertible non-increasing function such that*

$$\lim_{t\to-\infty} \psi(t) = +\infty, \ \psi(0) = 1, \ \textit{and} \ \lim_{t\to+\infty} \psi(t) = 0.$$

If ψ satisfies the condition (H_a) for some $a \in]0,1[$, then for all $s,t \in \mathbb{R}$,

$$\lim_{r\searrow 0} G_r(s,t) = 0 \Leftrightarrow \min(s,t) = 0.$$

Proof We prove the direct implication ($\Rightarrow$):

Let $s,t \in \mathbb{R}$ be fixed. By Lemma 3.1, for any $r > 0$ we have $G_r(s,t) \leq \min(s,t)$ and, then $\min(s,t) \geq 0$. We finish the proof by contradiction as follows. Assume that $s = \min(s,t) > 0$. Since ψ is nonincreasing and $s \leq t$, we have

$$\psi(s/r) + \psi(t/r) \leq 2\psi(s/r).$$

By assumption (H_a) and for r small enough, $2\psi(s/r) \leq \psi(as/r)$. Indeed, the ratio s/r goes to infinity as r goes to 0 because $s > 0$. Hence $\psi(s/r) + \psi(t/r) \leq \psi(as/r)$.

Now since ψ^{-1} is nonincreasing,

$$as/r \leq \psi^{-1}\left(\psi(s/r) + \psi(t/r)\right)$$

or equivalently, with r small enough, $\quad s \leq a^{-1}G_r(s,t)$.

Passing to the limit, $\lim_{r\searrow 0} G_r(s,t) = 0$ and, then $s \leq 0$ in contradiction with $s > 0$.

Now, we prove the converse ($\Leftarrow$):

Assume $s = \min(s,t)$ and (2) thus $s = 0$. Since $\psi(0) = 1$, we have

$$G_r(s,t) = r\psi^{-1}\left(1 + \psi(t/r)\right).$$

If $t = 0$ then $\lim_{r \searrow 0} G_r(s,t) = \lim_{r \searrow 0} r\psi^{-1}(2) = 0$.

If $t > 0$ then $\lim_{r \searrow 0} \psi(t/r) = 0$. Thus $\lim_{r \searrow 0} G_r(s,t) = 0$ by continuity of ψ^{-1}.

In both cases, we have $\lim_{r \searrow 0} G_r(s,t) = 0$. □

For both "extreme" examples, the assertion of Theorem 3.1 is clearly satisfied. Indeed, direct computations lead to

1. For $s > 0$ and $t > 0$ such that $\frac{1}{s} + \frac{1}{t} \leq \frac{1}{r}$, we have the following explicit expression

$$G_r^{(1)}(s,t) = \frac{st - r^2}{s + t + 2r}.$$

   Note that the denominator is not zero when s,t are non-negative even when $s = t = 0$. In addition, when $\min(s,t) > 0$ we have $\lim_{r \searrow 0} G_r^{(1)}(s,t) = \dfrac{st}{s+t} < \min(s,t)$.

2. For any $s,t \in \mathbb{R}$, we have the following explicit expresssion

$$G_r^{(2)}(s,t) = -r\log(e^{-s/r} + e^{-t/r}).$$

   Assume $s = \min(s,t)$. Then we have $s - r\log 2 \leq G_r^{(2)}(s,t)$ because

$$e^{-s/r} + e^{-t/r} \leq 2e^{-s/r}.$$

   Thus, $\quad \min(s,t) - r\log 2 \leq G_r^{(2)}(s,t) \leq \min(s,t)$.

   Passing to the limit as r goes to 0, we conclude that $\lim_{r \searrow 0} G_r^{(2)}(s,t) = \min(s,t)$.

Now, we focus on the case where ψ satisfies (H_a) for all $a \in]0,1[$ and prove a stronger result.

Theorem 3.2 *Let $\psi : \mathbb{R} \longrightarrow]0,+\infty[$ be an invertible non-increasing function such that*

$$\lim_{t \to -\infty} \psi(t) = +\infty,\ \psi(0) = 1 \text{ and } \lim_{t \to +\infty} \psi(t) = 0.$$

If ψ satisfies (H_a) for all $a \in]0,1[$, then for any $s,t > 0$,

$$\lim_{r \searrow 0} G_r(s,t) = \min(s,t).$$

Proof By Lemma 3.1, we have

$$\forall r > 0,\ \forall s,t \in \mathbb{R}, \quad \mathrm{G_r(s,t) \leq \min(s,t)}.$$

Thus, we have to concentrate on the lower bound of G_r.

Let $s,t > 0$ such that $s = \min(s,t) > 0$. For each $a \in]0,1[$ and when r is sufficiently small (i.e. $s/r \geq s_a > 0$) we can apply the assumption (H_a) to get

$$\psi(s/r) + \psi(t/r) \leq 2\psi(s/r) \leq \psi(as/r).$$

Since ψ^{-1} is nonincreasing, we deduce

$$as/r \leq \psi^{-1}\left(\psi(s/r) + \psi(t/r)\right).$$

Thus, for any $a \in]0,1[$ and any $0 < r < s/s_a$, we have $as \leq G_r(s,t)$.

Hence,

$$a \min(s,t) = as \leq \liminf_{r \searrow 0} G_r(s,t) \leq \limsup_{r \searrow 0} G_r(s,t) \leq \min(s,t).$$

By taking $a \nearrow 1$, we obtain the desired result . $\square$

Unfortunately, some functions (for instance $\theta^{(1)}$) don't satisfy the condition (H_a) for all $a \in]0,1[$. We have seen above that $\lim_{r \searrow 0} G_r^{(1)}(s,t) = \dfrac{st}{s+t}$, $(s,t > 0)$ which is strictly less than $\min(s,t)$.

Now our aim is to prove that $G_0(s,t) := \lim_{r \searrow 0} G_r(s,t)$ exists under some natural condition on ψ for fixed $s,t > 0$. For instance, the existence of such a limit is insured if the function $r \to G_r(s,t)$ is nonincreasing on some interval $]0,\varepsilon[$ (that is $\frac{\partial}{\partial r} G_r(s,t) \leq 0$ for $r \in]0,\varepsilon[$). We provide a sufficient and much more easier condition on ψ to fulfill this last technical condition.

Let V be a function defined in a positive neighborhood of 0. We say that V is locally subadditive at 0^+ if there exists $\eta > 0$ such that, for all $0 < \alpha,\ \beta,\ \alpha + \beta < \eta$, we have

$$V(\alpha + \beta) \leq V(\alpha) + V(\beta).$$

Theorem 3.3 *Let $\psi : \mathbb{R} \longrightarrow]0,+\infty[$ be a C^1 decreasing function such that*

$$\lim_{t\to-\infty} \psi(t) = +\infty, \psi(0) = 1 \textit{ and } \lim_{t\to+\infty} \psi(t) = 0.$$

(i) *Let $s,t > 0$ be fixed. If there exists $\varepsilon > 0$ such that $\frac{\partial}{\partial r} G_r(s,t) \leq 0$, $r \in]0,\varepsilon[$, then*

$$G_0(s,t) := \lim_{r\searrow 0} G_r(s,t) \text{ exists and } G_0(s,t) \leq \min(s,t).$$

(ii) *Suppose that $V := (-\psi' \circ \psi^{-1}) \times \psi^{-1}$ is locally subadditive at 0^+ then the conclusion of* (i)*holds true for any $s,t > 0$.*

(iii) *Let $s,t > 0$ be fixed. Assume that there exists $r_0 > 0$ such that for any $r \in]0,r_0]$,*

$$\frac{\partial}{\partial r} G_r(s,t) \leq 0 \textit{ and } r\frac{\partial}{\partial r} G_r(s,t) \leq G_r(s,t) - G_0(s,t),$$

then for any $r \in]0,r_0[$,

$$-r\,\frac{\big(G_0(s,t) - G_{r_0}(s,t)\big)}{r_0} + G_0(s,t) \leq G_r(s,t) \leq G_0(s,t). \tag{6}$$

Proof To simplify the presentation of the proof, we use the notation $f(r) := G_r(s,t)$ as a function of r when $s > 0$ and $t > 0$ are fixed.

(i) By Lemma 3.1, we always have $f(r) \leq \min(s,t)$. The assumption $f'(r) \leq 0$ implies that $f(r)$ is decreasing. Since f is bounded above by $\min(s,t)$,

$f(0^+) := \lim_{r\searrow 0} f(r)$ exists and satisfies $f(0^+) \leq \min(s,t)$.

(ii) We check the condition $\frac{\partial}{\partial r} G_r(s,t) = f'(r) \leq 0$ for any $s,t > 0$ of statement (i).

Step 1. We assume that $V := (-\psi' \circ \psi^{-1}) \times \psi^{-1}$ is locally subadditive at 0^+ that is that, there exists $\eta > 0$ such that for all $0 < \alpha, \beta, \alpha+\beta < \eta$, we have $V(\alpha+\beta) \leq V(\alpha) + V(\beta)$. Fix $s,t > 0$; since $\lim_{x\to+\infty} \psi(x) = 0$ and $\psi > 0$, there exists $\varepsilon > 0$ such that

$$0 < \max(\psi(s/\varepsilon), \psi(t/\varepsilon)) \leq \eta.$$

Thus for $r \in]0,\varepsilon[$, we have $0 < \alpha := \psi(s/r) < \eta$ and $0 < \beta := \psi(t/r) < \eta$ because ψ is decreasing.

By local subadditivity of V, we deduce $V(\alpha+\beta) \leq V(\alpha)+V(\beta)$ with α and β given above.

Step 2. A simple computation gives us

$$rf'(r) = f(r) - \frac{s\psi'(s/r)+t\psi'(t/r)}{(\psi'\circ\psi^{-1})(\psi(s/r)+\psi(t/r))}.$$

Thus the condition $f'(r) \leq 0$ is equivalent to

$$\psi^{-1}(\psi(s/r)+\psi(t/r)) \leq \frac{\frac{s}{r}\psi'(\frac{s}{r})+\frac{t}{r}\psi'(\frac{t}{r})}{(\psi'\circ\psi^{-1})(\psi(s/r)+\psi(t/r))}.$$

Let α and β as in step 1). We get

$$\psi^{-1}(\alpha+\beta) \leq \frac{(\psi'\circ\psi^{-1})(\alpha)\,\psi^{-1}(\alpha)+(\psi'\circ\psi^{-1})(\beta)\,\psi^{-1}(\beta)}{(\psi'\circ\psi^{-1})(\alpha+\beta)}.$$

Because $\psi' < 0$, this condition is exactly the subadditivity property of V i.e.

$$V(\alpha+\beta) \leq V(\alpha)+V(\beta). \tag{7}$$

Since the inequality (7) is satisfied, by Step 1 it implies that $f'(r) \leq 0$ for any $r \in (0,\varepsilon)$.

We apply (i) and conclude the proof of statement (ii) .

(iii) Now we prove (6). We have assumed that

$$rf'(r) \leq f(r) - f(0^+),\ 0 < r \leq r_0.$$

So,

$$\left(\frac{f(r)}{r}\right)' = \frac{rf'(r)-f(r)}{r^2} \leq \frac{-1}{r^2}f(0^+).$$

Let $0 < t < r_0$. By integration over the interval $[t, r_0]$ of the inequality just above, we get

$$\frac{f(r_0)}{r_0} - \frac{f(t)}{t} \leq f(0^+)\left(\frac{1}{r_0}-\frac{1}{t}\right),$$

and so, the desired result $\quad -t\left[\frac{f(0^+)-f(r_0)}{r_0}\right] + f(0^+) \leq f(t), \quad 0 < t \leq r_0.$ □

Remark 3.2 - *The assumption of subadditivity is a priori stronger than condition in* 1.

- *For fixed $s,t > 0$, the bounds on $G_r(s,t)$ in* (6) *give useful information for numerical* *simulation, since we have*

$$0 \le G_0(s,t) - G_r(s,t) \le \frac{r}{r_0}\left(G_0(s,t) - G_{r_0}(s,t)\right) \le \frac{r}{r_0}\left(\min(s,t) - G_{r_0}(s,t)\right).$$

Back to both "extreme" examples, we see much more easily that:

1. For $\psi^{(1)}(x) = \frac{1}{x+1}, x \ge 0$, $V^{(1)}(y) = y - y^2$, $0 < y < 1$ or $V^{(1)}(y) = 1 - y$, $y > 1$.

2. For $\psi^{(2)}(x) = e^{-x}, x \in \mathbb{R}$, $V^{(2)}(y) = -y \ln y$, $0 < y < \infty$.

The function $V^{(1)}$ and $V^{(2)}$ are clearly subadditive. A generic way to construct new functions ψ from old φ's such that the associated V_ψ is subadditive if V_φ is subadditive is as follows. Consider $\psi(x) = \varphi(\mu x^\lambda)$, $x \ge 0$ with nonnegative μ and ,λ, then we obtain $V_\psi = \lambda V_\varphi$. If V_φ is subadditive then V_ψ is subadditive. This applies to $x \mapsto (1+\mu x^\lambda)^{-1}$ and $x \mapsto e^{-\mu x^\lambda}$, $x \ge 0$, built on $\psi^{(1)}$ and $\psi^{(2)}$. Unfortunately V cannot be linear. Indeed, by solving the differential equation relating ψ and V it leads to $\psi(y) = \frac{c}{y^\gamma}$. But $\psi(0) = 1$ cannot be satisfied.

The condition 3) namely $rf'(r) \le f(r) - f(0)$ with $0 < r$ small enough is also satisfied for these two examples. Let $K = -\psi' \circ \psi^{-1}$. With the notations above, we have for $\alpha, \beta > 0$:

$$rf'(r) = f(r) - \frac{sK(\alpha) + tK(\beta)}{K(\alpha+\beta)}.$$

1. For $\psi^{(1)}$. We obtain

$$\frac{sK(\alpha)+tK(\beta)}{K(\alpha+\beta)} = s\left(\frac{\alpha}{\alpha+\beta}\right)^2 + t\left(\frac{\beta}{\alpha+\beta}\right)^2 \ge \inf_{0\le\lambda\le1}\{s\lambda^2 + t(1-\lambda)^2\} = \frac{st}{s+t} = f(0).$$

   Thus, we have $\quad rf'(r) \le f(r) - f(0)$.

2. For $\psi^{(2)}$. We obtain ($K = Id$),

$$\frac{sK(\alpha)+tK(\beta)}{K(\alpha+\beta)} = \frac{s\alpha + t\beta}{\alpha+\beta} \ge \min(s,t) = f(0).$$

   Thus $rf'(r) \le f(r) - f(0)$.

We remark that the two functions $G_r^{(1)}$ and $G_r^{(2)}$ and their limit function $G_0^{(1)}, G_0^{(2)}$ are concave on $(\mathbb{R}^2)^+$. This property is very important and can be used to obtain convergence results based on the smoothing technique discussed in [8]. The following theorem presents a simple and general necessary and sufficient condition on the smoothing functions to insure this property of concavity. First note that $G(s,t) = \psi^{-1}(\psi(s)+\psi(t))$ is a concave function of (s,t) with $s,t > 0$ if and only if $G_r(s,t) = rG(s/r,t/r)$ is a concave function for any $r > 0$ ($G = G_1$ with this notation).

Theorem 3.4 *Let* $\psi : \mathbb{R} \longrightarrow]0,+\infty[$ *be a* C^2 *non-increasing and strictly convex function* ($\psi' < 0$, $\psi'' > 0$) *and*

$$L(\alpha) := -\frac{(\psi' \circ \psi^{-1})^2}{\psi'' \circ \psi^{-1}}(\alpha), \quad \alpha \in \psi(\mathbb{R}).$$

The following statements are equivalent:

(i) *G is concave in the argument (s,t).*

(ii) *L is nonincreasing and subadditive i.e.* $L(\alpha+\beta) \le L(\alpha)+L(\beta), \quad \alpha,\beta \in \psi(\mathbb{R})$.

Moreover, if 1. *or* 2. *holds true and if* $G_0 = \lim_{r \searrow 0} G_r$ *exists then* G_0 *is concave.*

Proof To simplify the presentation of the proof, we denote $\alpha := \psi(s)$, $\beta := \psi(t)$,
$W := W(\alpha+\beta) = (\psi' \circ \psi^{-1})(\alpha+\beta) < 0$ (when $\alpha+\beta \in \psi(\mathbb{R})$ and
$U := U(\alpha+\beta) = (\psi'' \circ \psi^{-1})(\alpha+\beta) > 0$ (when $\alpha+\beta \in \psi(\mathbb{R})$.
A rather tedious computation leads to $R := \partial_{s,s}G(s,t) = \left[\psi''(s)W^2 - (\psi'(s))^2U\right]/W^3$,
$T := \partial_{t,t}G(s,t) = \left[\psi''(t)W^2 - (\psi'(t))^2U\right]/W^3$ and $S := \partial_{s,t}G(s,t) = -\psi'(s)\psi'(t)\frac{U}{W^3}$.
It is well-known that G is concave if and only if $R \le 0, T \le 0$ and $RT - S^2 \ge 0$. For the condition $R \le 0$ (similarly for $T \le 0$) and due to the fact that $W^3 < 0$, we get

$$\psi''(s)W^2 \ge (\psi'(s))^2U. \quad \text{or equivalently} \quad -\frac{(\psi'(s))^2}{\psi''(s)} \ge -\frac{W^2}{U}.$$

That is, with $s = \psi^{-1}(\alpha)$ and $t = \psi^{-1}(\beta)$, , that $L(\alpha) \ge L(\alpha+\beta)$, so that L is nonincreasing.

Now from the condition $RT - S^2 \ge 0$, we obtain after simplifications

$$W^6(RT - S^2) = \psi''(s)\,\psi''(t)W^4 - \left[\psi''(s)(\psi')^2(t) + \psi''(t)(\psi')^2(s)\right]W^2U.$$

By similar manipulations, as in the case of $R \le 0$, we can express the condition
$RT - S^2 \ge 0$ as the subadditivity of L. □

3.2 Construction of generic smoothing functions

The assumptions of Theorem 3.4 suggest a simple way to generate new smoothing functions. Indeed, an examination of both examples above leads us to the following important remark.

For the examples $\psi^{(1)}$ and $\psi^{(2)}$, we obtain respectively the following additive functions $L_1(\alpha) = -\frac{1}{2}\alpha$ and $L_2(\alpha) = -\alpha$. Then it suggests to construct a larger family of functions $\psi_{\lambda,c}$ including as particular cases our functions $\psi^{(1)}$ and $\psi^{(2)}$. This is done as follows. Since additivity property of L corresponds exactly to the equation $RT - S^2 = 0$. We can produce new functions ψ by solving this equation. It is easily seen that we can rewrite the equation $RT - S^2 = 0$ as $L(\alpha) = -\frac{1}{\lambda}\alpha$ for some $\lambda > 0$ (L is nonincreasing). Equivalently, we have $(\psi')^2 = \frac{1}{\lambda}\psi\psi''$.

First case: $\lambda > 1$. We obtain as solutions of this equation $\psi_{\lambda,c}(x) = \frac{1}{(cx+1)^{\frac{1}{\lambda-1}}}, \quad x > -c^{-1}$, for some constant $c > 0$.

Second case: $\lambda = 1$. The equation leads us to a different kind of family of solutions. Indeed, we get $\psi_{1,c}(x) = e^{-cx}$, $x \in \mathbb{R}$, for some constant $c > 0$.

In some sense, this instance is a limit case for the family of solutions of the first case when $\lambda \searrow 1$. Clearly both families of functions $(\psi_{1,c})_{c>0}$ and $(\psi_{\lambda,c})_{c>0,\lambda>1}$ have different behaviors. The first one has a polynomial decay and the second one has an exponential one. We note that $\psi_{2,1} = \psi^{(1)}$ and $\psi_{1,1} = \psi^{(2)}$. We also remark that all the smoothing functions generated this way satisfy (H_a). More precisely, the functions $\psi_{\lambda,1}$, when $\lambda > 1$ satisfy (H_a) with $a \in]0, (\frac{1}{2})^{\lambda-1}[$. The functions $\psi_{1,c}, c > 0$ satisfy (H_a) with $a \in]0,1[$.

All the smoothing functions generated this way satisfy the assumptions of Theorem 3.3 and satisfy $\psi_{\lambda,c}(x) \leq \psi^{(1)}$ when $1 < \lambda \leq 2$ and $c = 1$.

4 Convergence and Error Estimate

In this section, we propose a generic algorithm to solve (NCP) and prove some convergence results and error estimates. In what follows, when $r > 0$, we consider the function H_r defined by

$$H_r(x) := G_r(x, F(x)) = (G_r(x_i, F_i(x)))_{i=1}^n$$

where G_r defined by (5) and define $H_0(x) := \lim_{r\searrow 0} H_r(x) = \lim_{r\searrow 0} G_r(x, F(x))$ when the limit exists, for instance under the assumptions of Theorem 3.3.

The algorithm consists in finding the solution of a sequence of well-posed equations:

$$\left[\begin{array}{l} \text{Let } \{r^k\}_{k\in\mathbb{N}} \;/\quad r^0 > 0 \text{ and } \lim_{k\to\infty} r^k = 0, \\ \text{Find } x^k \;/\quad H_{r^k}(x^k) = 0. \end{array}\right.$$

The next lemma measures the "additional coercivity" effect of the smoothing.

Lemma 4.1 *Assume F is a P_0-function. then*

(i) *For any $r > 0$, H_r is a P-function.*

(ii) *If H_0 exists then it is a P_0-function.*

Proof (i) Let x, y be two distinct vectors of $\mathbb{R}^n$. Since F is a P_0-function there exits an index $i \in \{1, ..., n\}$ such that $x_i \neq y_i$ and $(x_i - y_i)(F_i(x) - F_i(y)) \geq 0$. Without loss of generality, we can suppose that $x_i > y_i$ and $F_i(x) \geq F_i(y)$.

Since ψ and ψ^{-1} are decreasing functions we obtain consecutively that for any $r > 0$,

$$\begin{aligned} &\psi(x_i/r) + \psi(F_i(x)/r) < \psi(y_i/r) + \psi(F_i(y)/r), \\ &G_r(x_i, F_i(x)) > G_r(y_i, F_i(y)). \end{aligned} \tag{8}$$

Hence, H_r is a P-function.

(ii) If H_0 exists, passing to the limit in (8) as $r \searrow 0$, we obtain that H_0 is a (P_0)-function. □

Using Theorem 3.3 and Lemma 4.1, we are now able to present a convergence result.

Theorem 4.1 *Assume that F is a P_0-function. Under the hypotheses of Theorem 3.3 on G_r, we have*

(i) *There exists an $\hat{r} > 0$ such that for any $0 < r < \hat{r}$, $H_r(x) = 0$ has a unique solution $x^{(r)}$ and the mapping $r \mapsto x^{(r)}$ is continuous on $(0, \hat{r})$.*

(ii) $\lim_{r\searrow 0} \mathrm{dist}(\mathrm{x}^{(\mathrm{r})}, \mathscr{Z}) = 0.$

Proof Using Theorem 3.3, the function G_0 exists so H_0 does. By Lemma 4.1, the function H_0 is a P_0-function and all functions H_r with $r > 0$ are P-functions. The functions H_r are continuous perturbations of H_0. This corresponds exactly to the assumptions of Theorem 4 (2) of [4]. So that (i) and (ii) are directly obtained by this theorem. $\square$

Remark 4.1 *Using the concavity results of Theorem* 3.4*, we can prove another convergence result based on the smoothing technique discussed in* [8].

When using $\psi \le \psi^{(1)}$ (this is the case of $\psi^{(2)}$ and of some functions $\psi_{\lambda,c}$), we can prove an estimate for the error term $||x^* - x^{(r)}||$ between the solution x^* and the approximation $x^{(r)}$ under a monotonicity assumption on F.

Proposition 4.1 *Assume that $\psi \le \psi^{(1)}$, x^* is a solution of (NCP) and $x^{(r)}, 0 < r < r_1$ is a sequence of non-negative solutions of $H_r(x) = 0$ for some $r_1 > 0$. Then*

(i) $x_i^{(r)} F_i(x^{(r)}) \le r^2$, $\forall i = 1 \dots n$.

(ii) *Furthermore, if F satisfies the condition*

$$h(||x-y||) \le \ (x-y, F(x)-F(y)) \tag{9}$$

with $h : [0,+\infty[\longrightarrow [0,+\infty[$ such that $h(0) = 0$, $h(t) > 0$ when $t > 0$ and there exist $\varepsilon, \eta > 0$ such that $h :]0,\varepsilon[\longrightarrow]0,\eta[$ is an increasing bijection. Then (NCP) has a unique solution namely x^* and there exists $r_0 > 0$ such that for any $r \in]0, r_0[$,

$$||x^* - x^{(r)}|| \le h^{-1}(nr^2). \tag{10}$$

Proof (i) Recall that $x^{(r)}$ satisfies $H_r(x^{(r)}) = 0$, i.e.

$$\psi\left(\frac{x_i^{(r)}}{r}\right) + \psi\left(\frac{F_i(x^{(r)})}{r}\right) = 1, \quad i : 1 \dots n.$$

Since $\psi \le \psi^{(1)}$, we obtain

$$\psi^{(1)}\left(\frac{x_i^{(r)}}{r}\right) + \psi^{(1)}\left(\frac{F_i(x^{(r)})}{r}\right) \ge 1, \quad i : 1 \dots n.$$

Then, a simple computation leads to $\qquad x_i^{(r)} F_i(x^{(r)}) \le r^2, \quad i : 1 \dots n.$

(ii) The uniqueness of x^* is a direct consequence of (9), and we have

$$\begin{aligned}(x^* - x^{(r)}, F(x^*) - F(x^{(r)})) &= (x^*, F(x^*) - F(x^{(r)})) - (x^{(r)}, F(x^*) - F(x^{(r)})) \\ &\leq (x^{(r)}, F(x^{(r)})) \leq nr^2.\end{aligned}$$

Indeed, $(x^*, F(x^*)) = 0$ and $-(x^*, F(x^{(r)})) - (x^{(r)}, F(x^*)) \leq 0$. By assumption (9), we immediately get

$$h(||x^* - x^{(r)}||) \leq nr^2.$$

Let r_0 such that $nr_0^2 < \eta$. Since h is a bijection from $[0, \varepsilon[$ onto $[0, \eta[$ and h^{-1} is increasing, we conclude (10) $\square$

5 Numerical Results

In this section, we present some numerical experiments for the two smoothing functions $\theta^{(1)}$ and $\theta^{(2)}$. Our aim is just to verify the theoretical assertions for these two ”extreme” cases. We consider ten test problems (that can be found in [9–14]) with various sizes and characteristics. Some of them are linear, the others nonlinear. In some cases, F is monotone or strongly monotone whereas others can have a non connected solution set; in this case F is at most a P_0-function.

A precise description of each test problem is given in the appendix. Since our method is fundamentally different from the existing methods, it is difficult to present any comparison. Nevertheless, we present in the appendix the numerical results obtained by the well-known projection iterative method (see [15, Section 12.1]) when it is used exactly in the same conditions. We used the following algorithm and heuristic updating strategy.

Algorithm.

Step 1. Let $x^0 > 0$, $\varepsilon > 0$ and set $\quad r^0 = \max\left(1, \sqrt{\max_{1\leq i\leq n} |x_i^0 F_i(x^0)|}\right).$

Step 2. If $\max\limits_{1\leq i\leq n} |x_i^k F_i(x^k)| \leq \varepsilon$ then stop.

Step 3. Compute x^k (an approximate solution of) $H_{r^k}(x^k) = 0$

(by using any Newton-type method).

Step 4. Update the parameter as follows $\quad r^{k+1} = \min\left(0.1r^k, (r^k)^2, \sqrt{\max\limits_{1\leq i\leq n} |x_i^k F_i(x^k)|}\right),$

and go back to **Step 2.**

We implemented this algorithm on a standard laptop (2.5 Ghz, 2Go M) in Matlab® and using the $fsolve$ function at Step 2. The stopping ε parameter is fixed to 10^{-8}.

We list in Table 1, the **worst** obtained results with respect to starting points. Indeed, for each test problem, we used **11** different starting points: a vector of ones, and 10 uniformly generated vectors with entries in $]0,20[$.

Pb	size	OutIter (θ_1,θ_2)	InIter (θ_1,θ_2)	Opt. (θ_1,θ_2)	Feas. (θ_1,θ_2)	cpu time (s) (θ_1,θ_2)
P1	10	$(6,4)$	$(65,15)$	$(5.6e–15,2.5e–18)$	$(1.1e–11,1.3e–10)$	$(0.22,0.09)$
	100	$(6,4)$	$(68,19)$	$(1.6e–14,7.1e–22)$	$(5.1e–13,1.4e–14)$	$(3.73,1.19)$
	500	$(6,4)$	$(83,21)$	$(5.4e–12,1.6e–16)$	$(1.9e–16,1.4e–14)$	$(31.15,89.26,)$
	1000	$(6,5)$	$(77,40)$	$(3.0e–14,3.1e–14)$	$(5.1e–18,1.8e–17)$	$(388.59,201.43)$
P2	10	$(6,4)$	$(79,23)$	$(2.1e–15,2.7e–15)$	$(7.6e–11,9.6e–19)$	$(0.31,0.11)$
	100	$(6,4)$	$(88,33)$	$(1.84e–12,1.0e–23)$	$(7.1e–10,3.1e–14)$	$(4.83,1.80)$
	500	$(6,4)$	$(96,41)$	$(6.5e–10,1.9e–16)$	$(6.6e–09,1.2e–12)$	$(112.14,49.59)$
	1000	$(6,5)$	$(114,67)$	$(1.0e–17,1.4e–23)$	$(2.4e–08,7.5e–18)$	$(530.42,328.15)$
P3	10	$(5,4)$	$(63,15)$	$(2.2e–12,2.7e–21)$	$(4.9e–08,1.4e–11)$	$(0.22,0.09)$
	100	$(5,4)$	$(71,18)$	$(7.9e–13,2.6e–15)$	$(9.5e–08,4.5e–08)$	$(3.10,1.02)$
	500	$(5,4)$	$(73,21)$	$(1.1e–14,2.6e–16)$	$(1.5e–07,5.9e–09)$	$(78.11,26.15)$
	1000	$(5,4)$	$(81,26)$	$(6.1e–13,1.2e–15)$	$(8.2e–10,2.4e–16)$	$(335.37,138.23)$
P4	4	$(6,4)$	$(63,20)$	$(5.4e–12,3.2e–17)$	$(6.1e–09,2.8e–12)$	$(0.15,0.08)$
P5	4	$(6,4)$	$(141,23)$	$(9.8e–14,2.1e–23)$	$(3.4e–07,3.2e–12)$	$(0.28,0.06)$
P6	5	$(5,3)$	$(47,17)$	$(1.3e–14,4.3e–27)$	$(4.9e–12,8.1e–17)$	$(0.16,0.07)$
P7	10	$(6,4)$	$(110,33)$	$(1.2e–16,6.1e–19)$	$(1.1e–12,4.5e–14)$	$(0.37,0.14)$
P8	20	$(6,5)$	$(145,66)$	$(2.9e–13,3.7e–21)$	$(0,4.4e–12)$	$(1.33,0.46)$
P9	30	$(6,6)$	$(106,77)$	$(3.7e–14,9.6e–21)$	$(4.4e–08,6.4e–11)$	$(2.24,0.85)$
P10	100	$(6,6)$	$(209,113)$	$(8.5e–11,2.1e–23)$	$(2.1e–07,1.8e–12)$	$(42.09,19.12)$

Table 1 Results for $\theta^{(1)}$ and $\theta^{(2)}$

In this table, **Size** stands for the number of variables, **OutIter** is the number of changes of the smoothing parameter, **InIter** corresponds to the total number of jacobian evaluations, **Opt.** and **Feas.** correspond to the following optimality and feasibility measures

Opt. $:= \max_{1\le i\le n} |x_iF_i(x)|$ and Feas. $:= \|\min(x,0)\|_1 + \|\min(F(x),0)\|_1$.

The results clearly show that our methods are efficient, competitive and superior to the projection one (The results of the projection method are given in the appendix). We also remark that the second smoothing function is much more efficient and powerful than the first one. This was foreseeable since $1-\delta_0(x) \geq \theta^{(2)}(x) \geq \theta^{(1)}(x)$, with $\delta_0(x) = 1$ if $x = 0$ and $\delta_0(x) = 0$ elsewhere. Of course other experiments with different θ's should be undertaken to assert the qualities of the whole approach.

6 Conclusions

We proposed simple methods to the solution of NCP and proved some convergence and error estimate results. We think that these methods can be much more improved in some special situations (Linear Complementarity Problems (LCP) for instance). We are going in this direction and our aim is to propose a polynomial algorithm for LCP with P-matrices. Further, for general NCP, additional work is needed to understand how best to choose among smoothing functions and control the decrease of the approximation parameter. Finally, additional computational testing and experimentation are needed to develop these algorithmic techniques into mature codes, which will thoroughly exploit the inherent characteristics of the smoothing strategy.

Acknowledgements The authors would like to thank anonymous referees and editors for their kind and helpful remarks and comments. The first author is partially supported by the ANR (Agence Nationale de la Recherche) through HJnet project ANR-12-BS01-0008-01.

7 Appendix

We give in this appendix a brief description of each test example and report some numerics obtained by using the following projection method; see [15, Sect. 12.1].

$$x^{k+1} = \max(0, x^k - D^{-1}F(x^k)), \quad k = 0, 1, \ldots$$

We choose $D = \lambda I$, where $\lambda > 0$ is a constant and I is the $n \times n$ identity matrix. Table 2 presents the best obtained results when varying the value of λ ($\lambda = 0.1, 1, 10, 20, 50, 100$).

-The two first examples **P1** and **P2** [9] correspond to strongly monotone function
$F(x) = (F_1(x), \ldots, F_n(x))^T$ with $F_i(x) = -x_{i+1} + 2x_i - x_{i-1} + \frac{1}{3}x_i^3 - b_i, \quad i = 1, \ldots, n,$
$(x_0 = x_{n+1} = 0)$ and $b_i = (-1)^i$ (resp. $b_i = \frac{(-1)^i}{\sqrt{i}}$), $\quad i = 1, \ldots, n$ for **P1** (resp. **P2**).
- **P3** is another strongly monotone test problem from [10] where $F(x) = (F_1(x), \ldots, F_n(x))^T$ with $F_i(x) = -x_{i+1} + 2x_i - x_{i-1} + \arctan(x_i) + (i - \frac{n}{2}),\ i = 1, \ldots, n,\ (x_0 = x_{n+1} = 0)$.
-**P4** and **P5** are known as the degenerate and the non-degenerate examples of Kojima-Shindo [12]. **P4** and **P5** are respectively defined by

$$F_4(x) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix}, F_5(x) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 3x_4 - 1 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix}.$$

P5 has a unique solution $x^* = (\frac{\sqrt{6}}{2}, 0, 0, \frac{1}{2})$ with $F(x^*) = (0, 2 + \frac{\sqrt{6}}{2}, 3, 0)$ while **P4** has two optimal solutions $x^* = (\frac{\sqrt{6}}{2}, 0, 0, \frac{1}{2})$ with $F(x^*) = (0, 2 + \frac{\sqrt{6}}{2}, 0, 0)$ and $x^{**} = (1, 0, 3, 0)$ with $F(x^{**}) = (0, 31, 0, 4)$. The first optimal solution of **P4** is degenerate since $x_3^* = F_3(x^*) = 0$.
-A complete description of **P6** and **P7** can be found in [14, 13]. These two examples correspond to the Nash-Cournot test problem with $N = 5$ and $N = 10$.
Let $x \in \mathbb{R}^{\mathrm{N}}$, $Q = \sum x_i$ and define the functions $C_i(x_i)$ and $p(Q)$ as follows:

$$p(Q) = 5000^{\frac{1}{\gamma}} Q^{\frac{-1}{\gamma}}, \quad C_i(x_i) = c_i x_i + \frac{b_i}{1 + b_i} L_i^{\frac{1}{b_i}} x_i^{\frac{b_i+1}{b_i}}.$$

The NCP-function is given by $F_i(x) = C_i{}'(x_i) - p(Q) - x_i p'(Q), \qquad i = 1, \ldots, N,$
or in a vectorial form $F(x) = \left[c + L^{\frac{1}{b}} x^{\frac{1}{b}} - p(Q)(e - \frac{x}{\gamma Q})\right]$
with c_i, L_i, b_i, $\gamma > 0$ and $\gamma \geq 1$.
For our numerics, we used - $N = 5$, $c = [10, 8, 6, 4, 2]^T$, $b = [1.2, 1.10, 1, 0.9, 0.8]^T$,
$L = [5, 5, 5, 5, 5]^T$, $e = [1, 1, 1, 1, 1]^T$ and $\gamma = 1.1$.
- $N = 10$, $c = [5, 3, 8, 5, 1, 3, 7, 4, 6, 3]^T$, $b = [1.2, 1, 0; 9, 0.6, 1.5, 1, 0.7, 1.1, 0.95, 0.75]^T$,
$L = [10, 10, 10, 10, 10, 10, 10, 10, 10, 10]^T$, $e = [1, 1, 1, 1, 1, 1, 1, 1, 1, 1]^T$ and $\gamma = 1.2$.
-**P8**, **P9** and **P10** are also described in [14, 13]. They correspond respectively to the HpHard test problem with $n = 20$, $n = 30$ and $n = 100$.

The corresponding function $F(x)$ is of the form: $F(x) = (AA^T + B + D)x + q$;

where the matrices A, B and D are randomly generated as: any entry of the square $n \times n$ matrix A and of the $n \times n$ skew-symmetric matrix B is uniformly generated from $]-5, 5[$, and any entry of the diagonal matrix D is uniformly generated from $]0, 3[$. The vector q is uniformly generated from $]-500; 0[$.

The matrix $AA^T + B + D$ is positive definite and the function F is strongly monotone. We used the M-files proposed in [13] to generate A, B, D and q. We implemented the projection method for solving the previous test problems in the same conditions and using the same material as for our methods. The following table gives the best obtained results when varying the value of λ ($\lambda = 0.1, 1, 10, 20, 50, 100$). In each computation we used a vector of ones as starting point. The column **Iter** corresponds to the number of iterations of the projection method and can not be compared to **Initer** or **Outiter** in Table 1. The other columns correspond to the same things in Table 1 and can be used for comparison.

Pb	size	Iter	cpu time (s)	Opt.
P1	10	71	0.63	$2.5e-9$
	100	72	5.48	$9.1e-11$
	500	89	96.37	$4.7e-10$
	1000	83	224.04	$8.8e-11$
P2	10	72	1.19	$2.2e-10$
	100	80	5.76	$7.1e-12$
	500	91	112.41	$5.3e-12$
	1000	102	336.20	$2.9e-11$
P3	10	41	1.03	$6.4e-11$
	100	73	5.19	$1.8e-12$
	500	82	90.22	$5.8e-13$
	1000	84	350.06	$2.4e-11$
P4	4	66	0.19	$3.1e-12$
P5	4	163	0.34	$1.4e-12$
P6	5	52	0.22	$6.5e-11$
P7	10	105	0.46	$2.7e-12$
P8	20	162	1.48	$9.3e-13$
P9	30	111	2.66	$1.6e-12$
P10	100	217	52.08	$8.7e-13$

Table 2 Results for the projection method

References

1. Complementarity: applications, algorithms and extensions. Papers from the International Conference on Complementarity (ICCP99) held in Madison, WI, June 9–12, 1999. Edited by Michael C. Ferris, Olvi L. Mangasarian and Jong-Shi Pang. Applied Optimization, 50. Kluwer Academic Publishers, Dordrecht (2001).
2. Ferris M. C., Pang J. S.: Engineering and economic applications of complementarity problems. SIAM Rev. 39, no. 4, 669–713 (1997).
3. Cottle R.W., Pang J.-S., Stone R.E.: The linear complementarity problem. Classics in Applied Mathematics 60. SIAM, Philadelphia, PA, USA (2009).
4. Seetharama Gowda M., Tawhid M. A.: Existence and limiting behavior of trajectories associated with P_0-equations. Computational optimization—a tribute to Olvi Mangasarian, Part I. Comput. Optim. Appl. 12, no. 1-3, 229–251 (1999).
5. Crouzeix J.-P.: Pseudomonotone variational inequality problems: Existence of solutions. Mathematical Programming, 78, 305–314 (1997).
6. Auslender A., Cominetti R., Haddou M.: Asymptotic analysis for penalty and barrier methods in convex and linear programming. Math. Oper. Res. 22, no. 1, 43–62 (1997).
7. Haddou M.: A new class of smoothing methods for mathematical programs with equilibrium constraints, Pacific Journal of Optimization, vol. 5(1), 86–96 (2009).
8. Ben-Tal A., Teboulle M.: A Smoothing Technique for Nondifferentiable Optimization Problems. In Dolecki, editor, Optimization, Lectures notes in Mathematics 1405, pages 1–11, New York. Springer Verlag (1989).
9. Huang C., Wang S.: A power penalty approach to a Nonlinear Complementarity Problem, Operations research letters, vol.38(1), 72–76 (2010).
10. Dong-hui Li, Jin-ping Zeng: A penalty technique for nonlineair problems, Journal of Computational Mathematics, vol.16(1), 40–50 (1998).
11. Ding Jundi, Yin Hongyou. : A new homotopy method for nonlinear complementarity problems. Numer. Math. J. Chin. Univ. (Engl. Ser.) 16, no. 2, 155–163 (2007).
12. Kojima M., Shindo S.: Extensions of Newton and quasi-Newton methods to systems of PC1 equations, J. Oper. Res. Soc. Jpn., 29, 352–374 (1986).
13. `http://dm.unife.it/pn2o/software/Extragradient/test_problems.html`
14. Harker P.T.: Accelerating the convergence of the diagonalization and projection algorithms for finite-dimensional variational inequalities, Mathematical Programming 48, 29–59 (1990).
15. Facchinei F., Pang J.S.: Finite-dimensional Variational Inequalities and Complementarity Problems, vol. I, II, in: Springer Series in Operations Research, Springer-Verlag, New York (2003).